Detecting Cocaine Use with Wearable Electrocardiogram Sensors

Annamalai Natarajan

Abhinav Parate, Gustavo Angarita, Edward Gaiser, Robert Malison, Benjamin Marlin, Deepak Ganesan

1 School of Computer Science, University of Massachusetts, Amherst
2 Department of Psychiatry, Yale School of Medicine, New Haven

Ubicomp 2013
September 10, 2013
Physiological Sensing and Addiction

- **Long-term goals**
  1. Improve our understanding of addiction
  2. Identify addiction triggers
  3. Design personalized interventions

- **In this paper:** We study the problem of detecting cocaine use based on physiological data collected from wearable on-body sensors
Cocaine - Facts

- Cocaine is a powerful, addictive stimulant drug made from coca plants native to South America.
- In 2009 there were 4.8 million cocaine users and 1 million crack cocaine users in the United States\(^1\).
- In 2012, global cocaine use was reported to be between 13.2 to 19.5 million users (adults aged 15-64)\(^2\).

---

\(^1\) National Survey on Drug Use and Health, 2009

\(^2\) World Drug Report 2012, United Nations Office on Drugs and Crime
Study Design

- National Institute on Drug Abuse (NIDA) approved study
- **Subjects**: Habituated, non-treatment seeking adults

*Cocaine Day*

1. **Session I**: Baseline (abbrev. B)
2. **Session II**: Fixed ascending dose regimen of 8mg, 16mg, 32mg (8, 16, 32)
3. **Session III**: Self administration sessions (A)

- All cocaine self-administration sessions take place at the Yale Center for Clinical Investigations Hospital Research Unit
Sample Heart Rate on Cocaine Day

![Graph showing heart rate changes over time with annotations B, 8, 16, 32, and SA.](image-url)
Cocaine - Short term physiological effects

[Schwartz et al., Tella et al., Foltin et al., Trippenbach et al., Regalado et al., Magnano et al.,
Levin et al., Hale et al.,]
Electrocardiogram (ECG) Morphology

- **Atrial Depolarisation**
- **Ventricular Depolarisation**
- **Ventricular Repolarisation**

- **PR Interval**
- **QRS Duration**
- **QT Interval**

- **Normal Heartbeat**
- **Fast Heartbeat**
- **Slow Heartbeat**
- **Irregular Heartbeat**

- **Activation of the atria**
- **Activation of the ventricles**
- **Recovery wave**
Physiological Sensing

**Zephyr BioHarness 3 chest band**
- 3 lead electrodes, no skin preparation
- ECG, breathing rate, 3 axis accelerometer, skin temperature
- 250 Hz, battery life ~12 hours, memory 500+ hours
- Smartphone app for instant viewing

**Smartphone**
- Samsung Nexus phones
- Chest band communicates to the phone via bluetooth
Data Collection

Subjects
- Analyzed data from six subjects

Behavior Data - Manually recorded
- Start and end times of sessions
- Dosage levels

ECG Packet - Zephyr BioHarness 3
- Raw ECG data every 4 milliseconds
Why is this problem hard?

1. Noise in data
Why is this problem hard?

2. Baseline shift
Why is this problem hard?

3. Sensor dropout
Detecting Cocaine Use with Wearable Electrocardiogram Sensors

Processing Pipeline

1. Data Logging
2. Peak Detection
3. Period Extraction
4. Local Averaging
5. Filtering
6. Sensing
Processed ECG periods

Smoothed Waveforms by Session

- Baseline
- 8mg
- 16mg
- 32mg
Feature Extraction

<table>
<thead>
<tr>
<th>Id</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>QTc</td>
<td>corrected distance between Q and T peaks</td>
</tr>
<tr>
<td>AM</td>
<td>All morphological features (QT, QTc, PR, QRS, TH)</td>
</tr>
<tr>
<td>W</td>
<td>Waveform Features</td>
</tr>
<tr>
<td>AM+W</td>
<td>All morphological plus waveform features</td>
</tr>
</tbody>
</table>

\[ QTc = \frac{QT}{\sqrt{RR}} \]  

\[^3\text{Bazett's correction, } QTc = \frac{QT}{\sqrt{RR}}\]
Within-Subject Classification

- Cocaine detection problem: Baseline vs. 8mg cocaine, Baseline vs. 16mg cocaine, etc
- Time preserved train and test set
- Linear logistic regression classifier

\[
p(y = 1|x, \beta) = \frac{1}{1 + \exp^{-(\beta_0 + x\beta)}} \tag{1}
\]

- Report the Area under Receiver Operating Characteristics curve (AUC) due to sample imbalances
Within-Subject Classification Results

The image shows a bar chart titled "Within Subjects: AUC vs Features". The x-axis represents different features such as RR, W, AM+W, AM, T, QT, QRS, PR, and QTc. The y-axis represents the AUC values. Each feature is represented by a different color, and the bars show the AUC values with error bars indicating variability or confidence intervals.
Between-Subjects Classification

- Six-fold cross validation
- Penalized logistic regression classifier (penalty $= \lambda$)
- Choose hyper parameter, $\lambda$, based on training data
- Logistic regression $\beta$’s fitted using minFunc\(^4\) toolbox

\(^4\) http://www.di.ens.fr/~mschmidt/Software/minFunc.html
Between-Subjects Classification Results

Between Subjects: AUC vs Features

W AM+W AM0.0
0.2
0.4
0.6
0.8
1.0 AUC

Bv8
Bv16
Bv32
BvA
Conclusion

1. Collected wireless ECG data from experienced cocaine users in clinical settings
2. Developed a computational pipeline for inferring morphological features from noisy ECG waveforms
3. Reliably detect cocaine use based on data from wearable ECG sensors
4. Waveform features (data-driven) to cocaine detection is as effective as morphological features (knowledge-based)
Future Work

- **More Sensors**: Experimenting with new wrist band sensors in these settings
- **Better Models**:
  - Probabilistic model to simultaneously label all peaks in raw ECG data
  - Experiment with non-linear classifiers
- **Data Analysis**: Use click and infusion data to study craving attacks
- **Real world Deployment**: Methods to deploy this system outside of the clinical settings
Acknowledgements

- Clinical Neuroscience Research Unit at the Connecticut Mental Health Center and Hospital Research Unit at Yale-New Haven Hospital
- Department of Mental Health and Addiction Services of the State of Connecticut
- National Science Foundation
- National Center for Research Resources
- National Center for Advancing Translational Science
- President’s Science and Technology fund, University of Massachusetts, Amherst